ORIGINAL ARTICLE

Modeling Neurocognitive Decline and Recovery During Repeated Cycles of Extended Sleep and Chronic Sleep Deficiency

Melissa A. St. Hilaire, PhD[1,2,*]; Melanie Rüger, PhD[1,2,*]; Federico Fratelli, BSc[1–3]; Joseph T. Hull, PhD[1,2]; Andrew J. K. Phillips, PhD[1,2]; Steven W. Lockley, PhD[1,2]

[1]Division of Sleep and Circadian Disorders, Departments of Medicine and Neurology, Brigham and Women's Hospital, Boston, MA; [2]Department of Medicine, Division of Sleep Medicine, Harvard Medical School, Boston, MA; [3]Department of Medicine, Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom

*The first two authors are equal contributors to this work.

Study Objectives: Intraindividual night-to-night sleep duration is often insufficient and variable. Here we report the effects of such chronic variable sleep deficiency on neurobehavioral performance and the ability of state-of-the-art models to predict these changes.

Methods: Eight healthy males (mean age ± *SD*: 23.9 ± 2.4 years) studied at our inpatient intensive physiologic monitoring unit completed an 11-day protocol with a baseline 10-hour sleep opportunity and three cycles of two 3-hour time-in-bed (TIB) and one 10-hour TIB sleep opportunities. Participants received one of three polychromatic white light interventions (200 lux 4100K, 200 or 400 lux 17000K) for 3.5 hours on the morning following the second 3-hour TIB opportunity each cycle. Neurocognitive performance was assessed using the psychomotor vigilance test (PVT) administered every 1–2 hours. PVT data were compared to predictions of five group-average mathematical models that incorporate chronic sleep loss functions.

Results: While PVT performance deteriorated cumulatively following each cycle of two 3-hour sleep opportunities, and improved following each 10-hour sleep opportunity, performance declined cumulatively throughout the protocol at a more accelerated rate than predicted by state-of-the-art group-average mathematical models. Subjective sleepiness did not reflect performance. The light interventions had minimal effect.

Conclusions: Despite apparent recovery following each extended sleep opportunity, residual performance impairment remained and deteriorated rapidly when rechallenged with subsequent sleep loss. None of the group-average models were capable of predicting *both* the build-up in impairment and recovery profile of performance observed at the group or individual level, raising concerns regarding their use in real-world settings to predict performance and improve safety.

Keywords: chronic variable sleep deficiency, neurobehavioral performance, subjective sleepiness, recovery sleep, physiological adaptation, recovery of function.

Statement of Significance

Chronic sleep deficiency is endemic in modern society. Our findings not only confirm that individuals remain subjectively unaware of the impact of chronic sleep deficiency on their neurobehavioral function but also reveal that intermittent nights of extended sleep do not permanently restore this function despite apparent recovery. Additionally, existing group-average mathematical models that have been developed and are in use as fatigue management strategies to anticipate the effects of chronic sleep deficiency on neurobehavioral performance fail to predict our findings. Future work is therefore needed both to further our understanding of the mechanisms underlying recovery sleep and to refine existing group-average mathematical models to account for the variable sleep–wake patterns that are experienced in the real world.

INTRODUCTION

Sleep is an essential behavior that many individuals either consciously sacrifice to satisfy the economic and societal pressures of their lifestyles or are required to forgo, for example due to childcare or employment demands. According to a recent National Sleep Foundation poll, Americans obtain an average of less than 7 hours of sleep on weeknights, which two-thirds of people report to be insufficient.[1] A similar national survey in Great Britain found that 47% of the population obtains less than 7 hours of sleep on weeknights, and only 13% obtain 8 hours or more.[2] Studies have shown that less than 7 hours of sleep per night is associated with impaired reaction time and increased reports of daytime sleepiness.[3]

Human sleep is regulated by two distinct processes: the sleep–wake homeostat (Process S) and the circadian timing system (Process C).[4] Process S represents the build-up of sleep pressure as a function of time awake, which is dissipated by sleep, and regulates the amount of slow-wave sleep (SWS).[5] Process C generates an intrinsic near-24-hour rhythm in sleep propensity that is entrained to 24 hours primarily by the light–dark cycle and is typically maximal during the second half of the night. If the amount of sleep obtained during the night is insufficient to completely dissipate accumulated sleep pressure, then sleep "debt" remains and accumulates over time, reflected

by the continual deterioration of neurocognitive performance with increasing severity of chronic stable sleep deficiency.[3,6] Under real-world conditions, however, a more variable sleep pattern is often observed, with long recovery sleep episodes interspersed with sleep restriction.[7–10]

A common misconception is that the negative effects of chronic sleep deficiency can be dispelled by one to two nights of recovery sleep, for example over a weekend.[11] Although one extended sleep opportunity (10 hours) can have apparently substantial recovery effects,[12] it cannot fully restore performance to baseline (BL) levels. Even up to seven recovery nights (8-hour sleep opportunities) may be insufficient to recover performance, depending on the extent of the chronic stable sleep deficiency and pre-study sleep history.[3,12–15]

Borbély's seminal two-process model,[4] which describes the interaction between Process S and Process C, has provided the foundation for several models of the effects of sleep and circadian rhythms on neurobehavioral performance and subjective alertness (as reviewed in Refs[16–18]). Although these models are capable of capturing neurobehavioral performance and subjective alertness dynamics in response to up to 72 hours of continuous wakefulness, the models are unable to reproduce the effects of chronic sleep restriction. Recently, several models[19–21] have been introduced that

extend the two-process model to account for neurobehavioral performance and subjective alertness in response to both acute and chronic sleep loss. In addition, a physiologically based model of the adenosine system, which has been identified as a key physiological pathway mediating the effects of acute and chronic sleep restriction on sleep and performance, has been proposed (Supplementary Materials). These models fare reasonably well against data from laboratory-based studies of chronic sleep restriction and recovery with stable sleep opportunities day to day, but have not yet been rigorously tested under other scenarios.

No prior studies have examined the effects of chronic variable sleep deficiency (CVSD) on neurocognitive performance and subjective sleepiness, which may be more representative of real-world experiences, nor have they assessed whether the resultant performance patterns can be predicted by currently available models. We therefore investigated the effects of three cycles of two 3-hour time-in-bed (TIB) opportunities, interspersed with one 10-hour TIB opportunity, on neurocognitive performance and sleepiness. We also simulated the experimental protocol with five state-of-the-art group-average mathematical models to compare model predictions with data.

METHODS

Participants

Twelve male participants (mean age ± *SD*: 23.8 ± 2.2 years) were admitted for an 11-day study in an environment free of time cues in the Intensive Physiological Monitoring Unit of the Center for Clinical Investigation at the Brigham and Women's Hospital, Boston, MA, between September 2009 and February 2010. Participants were screened for medical, ophthalmological, and psychological health via examination, questionnaires, interview, and comprehensive urine and blood tests. Participants were instructed to maintain a fixed 8-hour nightly sleep schedule for two weeks, followed by at least 5 days during which participants moved their habitual sleep time 2 hours earlier for a total of 10 hours sleep opportunity per night prior to laboratory admission. Compliance was verified by calls to a time- and date-stamped voicemail, in addition to actigraphy for at least 1 week prior to admission (Actiwatch-L, Philips-Respironics, The Netherlands). Participants were required to abstain from caffeine, nicotine, alcohol, and other foreign substances from the beginning of screening until completion of the study; compliance was evaluated by toxicology tests during screening and upon admission. The study was approved by the Partners Human Research Committee in compliance with the Declaration of Helsinki. All participants gave written informed consent and were paid for their participation.

Experimental Protocol

During the 11-day protocol, participants remained in an individual time-free suite. The protocol started with a baseline (BL) day of 10 hours scheduled sleep (solid black boxes, Figure 1) and 14 hours scheduled wake (white boxes, Figure 1). The clock times of the BL sleep and wake episodes were based on the average of the self-reported sleep–wake call-in times from the 7 days prior to admission. The first cycle of sleep restriction started on day 2, consisting of two 3-hour TIB opportunities (SR1–SR2), followed by one 10-hour TIB opportunity (REC1), aligned according to the scheduled habitual wake time. This pattern was repeated twice more. Participants received breakfast, lunch, dinner, and a snack at 2.75, 6.75, 11.75, and 14.75 hours after wake time, respectively, during each day of the protocol. Participants were scheduled to four 10-hour constant postures (CPs; BL, SR2, SR4, and SR6; black striped boxes, Figure 1), during which they remained awake in bed in a semi-recumbent posture. During each CP, participants received a 3.5-hour light exposure (bolded black boxes, Figure 1). The same light exposure was administered during 180-minute sleep inertia test batteries (SIT; dashed gray boxes, Figure 1) completed on the mornings following REC1, REC2, and REC3 (same light as CP on SR2, SR4, and SR6, respectively).

Light Conditions

Ambient light was provided by 4100K fluorescent lamps (Philips Lighting, The Netherlands) with digital ballasts (Lutron Electronics Co., Inc, PA) transmitted through a UV-stable filter (Lexan 9030 with prismatic lens, GE Plastics, MA) and light levels were approximately 64 µW/cm^2 (~200 lux) at the level of the eye in the approximate angle of gaze for the first 14 hours of wake. Ambient light was lowered during the last 7 hours of wake (preceding each 3-hour TIB opportunity only) to approximately 23 µW/cm^2 (~89 lux) at 137 cm from the floor in the vertical plane, with a maximum level of 48 µW/cm^2 (~150 lux) when measured in the horizontal plane at a height of 187 cm from anywhere in the room.

The variable sleep schedule was initially designed to generate sufficient daytime sleepiness during waking hours to test the efficacy of blue-enriched white light compared with standard white light to improve daytime alertness and performance, given the short-wavelength sensitivity to the alerting effects of light.[22] The details of the experimental light exposures are outlined in Supplementary Materials.

Neurobehavioral Performance

Participants completed cognitive test batteries every 1–2 hours (every 1 hour for the first 9 hours and every 2 hours thereafter for BL and SR1–SR6; every 2 hours for REC1 and REC2) while awake throughout the study, starting 15 minutes after wake time. Here we report the results for the 10-minute version of the psychomotor vigilance test (PVT) and the Karolinska Sleepiness Scale (KSS). Further details about the cognitive test batteries, including results from the Addition calculation performance task (ADD) and the Unstable Tracking task (TRACK), are provided in Supplementary Materials.

Polysomnography Data

All polysomnography data were scored according to the Rechtschaffen and Kales criteria.[23] The number of minutes of Stage 1, 2, SWS, rapid eye movement sleep (REM), and total sleep time (TST) for each sleep episode were averaged across participants (Table 1).
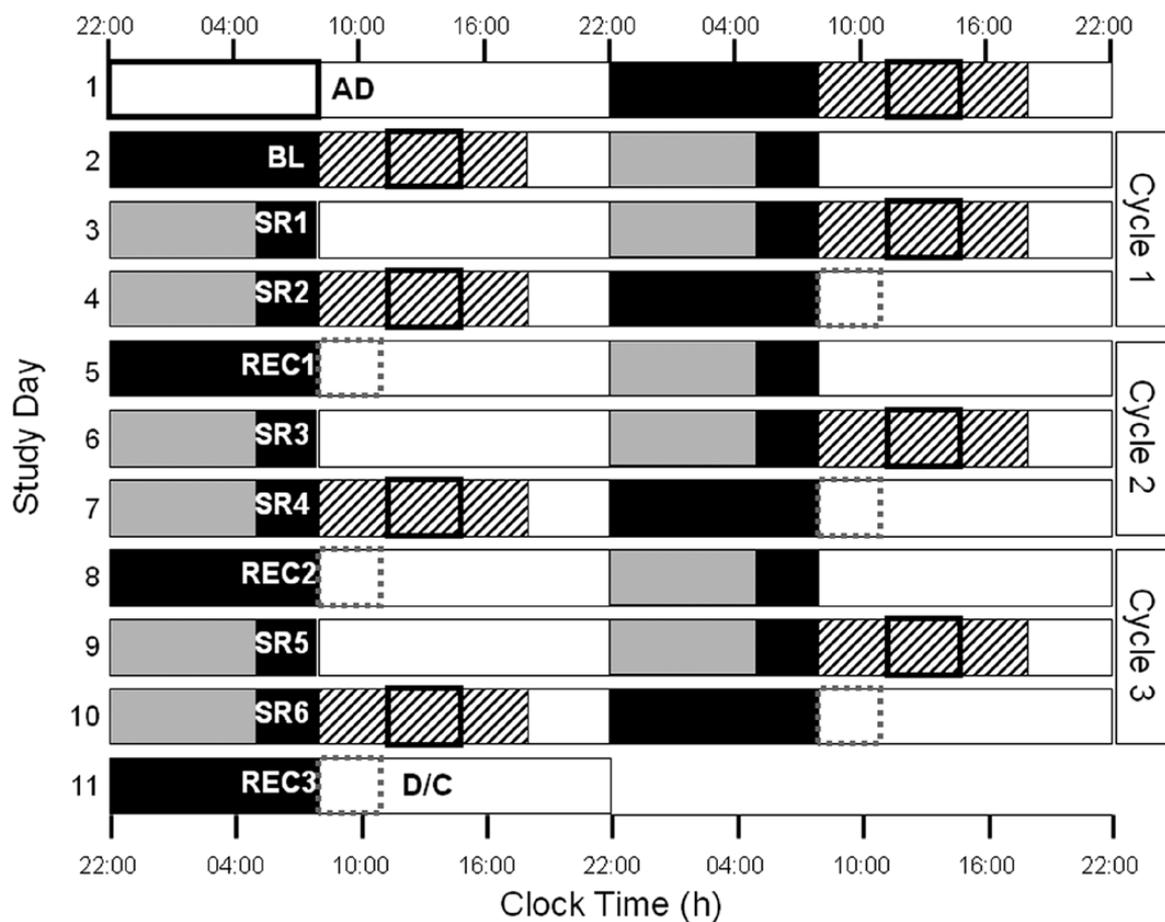
Figure 1—An 11-day inpatient chronic variable sleep deficiency protocol. The *y*-axis depicts consecutive study days and the *x*-axis depicts relative clock time. White boxes indicate wake episodes in ambient light of <200 lux (white bars) or <90 lux (gray bars), respectively; black bars indicate scheduled sleep in 0 lux. Black striped boxes indicate 10-hour constant postures (CP; BL, SR2, SR4, and SR6); bolded black boxes within the striped boxes indicate the 3.5-hour light exposures (200 lux 4100K on BL and SR2, randomized 200 or 400 lux 17000K on SR4 and SR6, respectively). Dashed gray boxes (REC1, REC2, and REC3) indicate sleep inertia testing during CP. AD = Admit, D/C = Discharge.

Data Analysis

For all analyses and figures, BL refers to the second wake period of the protocol following the first 10-hour sleep opportunity in the laboratory, which was preceded by approximately 1 week of 10 hours TIB opportunity at home. The sleep restriction days (SR1–6) refer to days following 3-hour TIB sleep opportunities and the recovery days (REC1–2) refer to days following 10-hour TIB sleep opportunities.

We analyzed the PVT (reciprocal mean reaction time [RT], fastest 10% mean RT, and number of lapses) and KSS for test batteries common to all study days outside of the light exposures, that is, test batteries occurring 8 and 10 hours after waketime. The PVT and KSS were analyzed using a repeated measures ANOVA in PROC MIXED (SAS 9.3, SAS, Inc, Cary, NC) for the within-subjects effects of cycle (cycle 1 = BL, SR1, SR2; cycle 2 = REC1, SR3, SR4; cycle 3 = REC2, SR5, SR6), condition (10-hour TIB = BL, REC1, REC2; first 3-hour TIB = SR1, SR3, SR5; second 3-hour TIB = SR2, SR4, SR6), and time (8 hours after wake and 10 hours after wake). A two-way ANOVA was conducted in PROC MIXED to test the effect of each 10-hour TIB recovery opportunity (BL, REC1,

and REC2) and time on each performance measure. A one-way ANOVA was also conducted in PROC MIXED to test differences in sleep stages (TST, S1, S2, SWS, and REM) during each 10-hour TIB opportunity. Post hoc pairwise comparisons (PROC MULTTEST, SAS 9.3) were adjusted for multiple comparisons using a Bonferroni correction on $\alpha \leq 0.05$. We computed effect sizes (d_z) for our post hoc comparisons in G*Power v3.1.9.2 (Düsseldorf, Germany) assuming a correlation between groups of 0.50.

Mathematical Modeling

We conducted an out-of-the-box challenge of several state-of-the-art group-average mathematical models. We simulated the 11-day CVSD protocol with three group-average mathematical models based on the two-process model of sleep regulation: the two-process model,[4,24] the state-space model,[19] and the unified model.[21] We also simulated this protocol with a new physiologically based model that attempts to link cognitive performance to underlying mechanistic pathways. Specifically, the model simulates the dynamics of cerebral extracellular adenosine concentration and A_1 adenosine

Table 1—Descriptive Statistics (Mean ± *SD*) of the Amount of Sleep Across Sleep Episodes.

Day	TST	S1		S2		SWS		REM	
	Min	Min	%	Min	%	Min	%	Min	%
	Mean ± *SD*	Mean ± *SD*	Meanw ± *SD*	Mean ± *SD*	Mean ± *SD*	Mean ± *SD*	Mean ± *SD*	Mean ± *SD*	Mean ± *SD*
BL	541.9 (± 51.6)	70.8 (± 23.5)	13.0 (± 3.8)	285.1 (± 32.7)	52.6 (± 3.2)	81.6 (± 12.2)	15.2 (± 2.5)	104.4 (± 23.1)	19.2 (± 3.5)
SR1	172.9 (± 3.7)	16.4 (± 12.1)	9.5 (± 7.0)	60.7 (± 13.6)	35.1 (± 7.7)	64.2 (± 14.0)	37.2 (± 8.1)	31.6 (± 6.1)	18.3 (± 3.5)
SR2	176.1 (± 2.6)	10.6 (± 7.8)	6.0 (± 4.5)	57.1 (± 16.1)	32.4 (± 9.2)	76.2 (± 18.5)	43.3 (± 10.4)	32.3 (± 10.0)	18.3 (± 5.7)
REC1	569.5 (± 30.5)	50.6 (± 24.7)	8.9 (± 4.2)	274.2 (± 46.6)	48.0 (± 7.0)	114.9 (± 22.6)	20.3 (± 4.5)	129.8 (± 22.2)	22.8 (± 3.8)
SR3	170.6 (± 7.8)	13.8 (± 6.4)	8.1 (± 4.0)	50.2 (± 17.0)	29.3 (± 9.5)	74.9 (± 15.3)	44.1 (± 10.1)	31.8 (± 13.7)	18.5 (± 7.5)
SR4	176.9 (± 3.7)	11.7 (± 7.0)	6.7 (± 4.1)	44.1 (± 14.8)	24.9 (± 8.2)	85.8 (± 18.0)	48.5 (± 10.0)	35.3 (± 7.5)	20.0 (± 4.3)
REC2	575.1 (± 10.8)	44.9 (± 17.6)	7.8 (± 3.0)	280.3 (± 38.1)	48.8 (± 7.3)	109.9 (± 19.2)	19.1 (± 3.2)	139.9 (± 27.6)	24.3 (± 4.6)
SR5	174.9 (± 4.9)	11.4 (± 3.6)	6.5 (± 2.1)	54.8 (± 16.3)	31.2 (± 8.9)	72.2 (± 11.3)	41.4 (± 6.8)	36.6 (± 7.1)	20.9 (± 4.0)
SR6	172.3 (± 8.6)	14.2 (± 7.3)	8.2 (± 4.1)	56.5 (± 23.1)	33.1 (± 14.3)	67.4 (± 20.7)	38.7 (± 11.4)	34.2 (± 15.8)	19.9 (± 9.1)

TST = total sleep time; SWS = slow-wave sleep; REM = rapid eye movement sleep; BL = baseline.

receptor concentration, and links these to PVT performance (see Supplementary Materials for equations). In addition, we simulated this protocol with a more recent version of the state-space model that incorporates a time dependence in the amplitude of the circadian modulation of performance.[20] The details of the equations for each published model can be found in the original publications. All models were coded and simulated in MATLAB v2014b (The MathWorks, Inc, Natick, MA). A complete list of parameters used for the adenosine model (which were derived by fitting the model to the data published in McCauley et al.[19]) can be found in Supplementary Table 1, and a complete list of parameters used for the published models (two-process model, state-space model, modified state-space model, and unified model) can be found in Supplementary Table 2. The differential equations were solved using a fourth-order Runge–Kutta method. Given that sleep efficiency was high across all sleep episodes, scheduled sleep amounts rather than actual sleep durations were used in the simulations. Simulations of the 11-day CVSD protocol were preceded by 1 week of habitual sleep (10 hours of sleep and 14 hours of wake) to stabilize model output prior to the predictions of interest.

The goodness-of-fit (GoF) of each model to the population-average PVT lapse data from the CVSD protocol was calculated using the root mean square error (RMSE), Spearman's rho, adjusted R^2, and Akaike's information criterion (AIC), which was computed as

$$\mathrm{AIC} = 2k + n \ln \frac{\mathrm{RSS}}{n}$$

where k is the number of parameters of the model (as determined from Supplementary Table 1 for the adenosine model and Supplementary Table 2 for the two-process, state-space, modified state-space, and unified models), n is the number of data points, and RSS is the residual sum of squares. We excluded data points representing PVT lapses recorded during

the first 3 hours of each wake episode, as none of the models tested here include a sleep inertia component.

RESULTS

Participants
Twelve participants were initially enrolled. Three participants withdrew due to excessive sleepiness on SR2, and a fourth participant withdrew for personal reasons on REC2. The remaining eight participants (mean age ± *SD*: 23.9 ± 2.4 years) were included in the analysis.

Effects of Experimental Light Exposure
There were minimal differences in performance between the experimental light exposures performed during BL, SR2, SR4, and SR6 (Supplementary Materials). Given the negligible impact of light on performance and subjective sleepiness observed, we analyzed the role of CVSD on performance and subjective sleepiness across all performance tests, including those during light exposure.

Sleep Parameters
There were no significant differences in TST across the three 10-hour TIB opportunities ($F_{(2,14)} = 3.01$, $p = .082$) (Table 1). The amount of S1 significantly differed across the three 10-hour TIB opportunities ($F_{(2,14)} = 17.71$, $p < .0001$) with less S1 during REC2 compared with BL ($p = .029$) but no difference in S1 between BL and REC1 ($p = .083$) or REC1 and REC2 ($p = .609$). There was no significant difference in S2 amounts across the three 10-hour TIB opportunities ($F_{(2,14)} = 0.43$, $p = .661$). The amount of SWS significantly differed across the three 10-hour TIB opportunities ($F_{(2,14)} = 12.76$, $p = .0007$) with more SWS during REC1 and REC2 compared with BL ($p = .002$ and $p = .006$, respectively) but no difference in SWS between REC1 and REC2 ($p = .600$). The amount of REM sleep significantly differed across the three 10-hour TIB opportunities ($F_{(2,14)} = 5.78$, $p = .015$) with more REM during REC1 and

REC2 compared with BL ($p = .049$ and $p = .008$, respectively) but no difference in REM between REC1 and REC2 ($p = .420$).

Main Effects of Condition

There was a significant main effect of condition on the PVT and KSS 8 and 10 hours after wake: reciprocal mean RT ($F_{(2,14)} = 39.85$, $p < .0001$), fastest 10% mean RT ($F_{(2,14)} = 26.22$, $p < .0001$, number of lapses ($F_{(2,14)} = 17.85$, $p = .0001$), and subjective sleepiness ($F_{(2,12)} = 18.80$, $p = .0001$). Post hoc pairwise comparisons showed that the reciprocal mean RT following 10 hours TIB differed significantly from the first and second 3-hour TIB opportunities ($p < .0001$ for both comparisons; $d_z = 1.31$ and $d_z = 1.48$, respectively); the reciprocal mean RT also differed significantly between the two 3-hour TIB opportunities ($p = .0369$, $d_z = 0.39$). This same pattern was observed for the number of lapses (10-hour TIB vs. first 3-hour TIB, $p = .0001$, $d_z = 0.86$; 10-hour TIB vs. second 3-hour TIB, $p < .0001$, $d_z = 1.13$; first 3-hour TIB vs. second 3-hour TIB, $p = .0037$, $d_z = 0.57$) and the KSS (10-hour TIB vs. first 3-hour TIB, $p < .0001$, $d_z = 0.97$; 10-hour TIB vs. second 3-hour TIB, $p < .0001$, $d_z = 1.35$; first 3-hour TIB vs. second 3-hour TIB, $p = .0447$, $d_z = 0.46$). For the fastest 10% mean RT, the 10-hour TIB opportunity differed significantly from the first and second 3-hour TIB opportunities ($p < .0001$ for both comparisons, $d_z = 1.05$ and $d_z = 1.13$, respectively), but the difference between the two 3-hour TIB opportunities ($p = .5664$, $d_z = 0.08$) did not reach significance.

Main Effects of Cycle

There was a significant difference across cycles (cycle 1 vs. cycle 2 vs. cycle 3) for the reciprocal mean RT ($F_{(2,14)} = 15.61$, $p = .0003$), fastest 10% mean RT ($F_{(2,14)} = 7.29$, $p = .0068$), and number of lapses ($F_{(2,14)} = 10.15$, $p = .0019$) (Figure 2A–C, respectively) 8 and 10 hours after waking, with performance declining over the three cycles of sleep deficiency. Post hoc pairwise comparisons showed that changes in performance were not consistent between cycles for all measures. The difference was significant between cycles 1 and 3 and cycles 2 and 3, but not significant between cycles 1 and 2, for the reciprocal mean RT (cycle 1–2: $p = .0696$, $d_z = 0.44$; cycle 2–3: $p = .0258$, $d_z = 0.46$; and cycle 1–3: $p < .0001$, $d_z = 0.88$) and the fastest 10% mean RT (cycle 1–2: $p = .3449$, $d_z = 0.33$; cycle 2–3: $p = .0325$, $d_z = 0.40$; and cycle 1–3: $p = .0022$, $d_z = 0.72$). For lapses, however, a significant difference was observed between cycles 1 and 3 only (cycle 1–2: $p = .0602$, $d_z = 0.41$; cycle 2–3: $p = .1316$, $d_z = 0.35$; and cycle 1–3: $p = .0008$, $d_z = 0.67$). There was no significant main effect of cycle on KSS ($F_{(2,14)} = 0.23$, $p = .7995$, Figure 2D).

Main Effect of Recovery Sleep

The reciprocal mean RT ($F_{(2,14)} = 17.81$, $p = .0001$), the fastest 10% mean RT ($F_{(2,14)} = 7.58$, $p = .0059$), and the number of lapses ($F_{(2,14)} = 7.83$, $p = .0052$) 8 and 10 hours after waking differed significantly following each of the three 10-hour TIB opportunities. Post hoc pairwise comparisons showed a significant difference in the reciprocal mean RT between BL and REC2 ($p = .003$, $d_z = 1.06$) and between REC1 and REC2 ($p = .012$, $d_z = 0.85$) with longer RTs observed following REC2. The fastest 10% mean RT was significantly higher
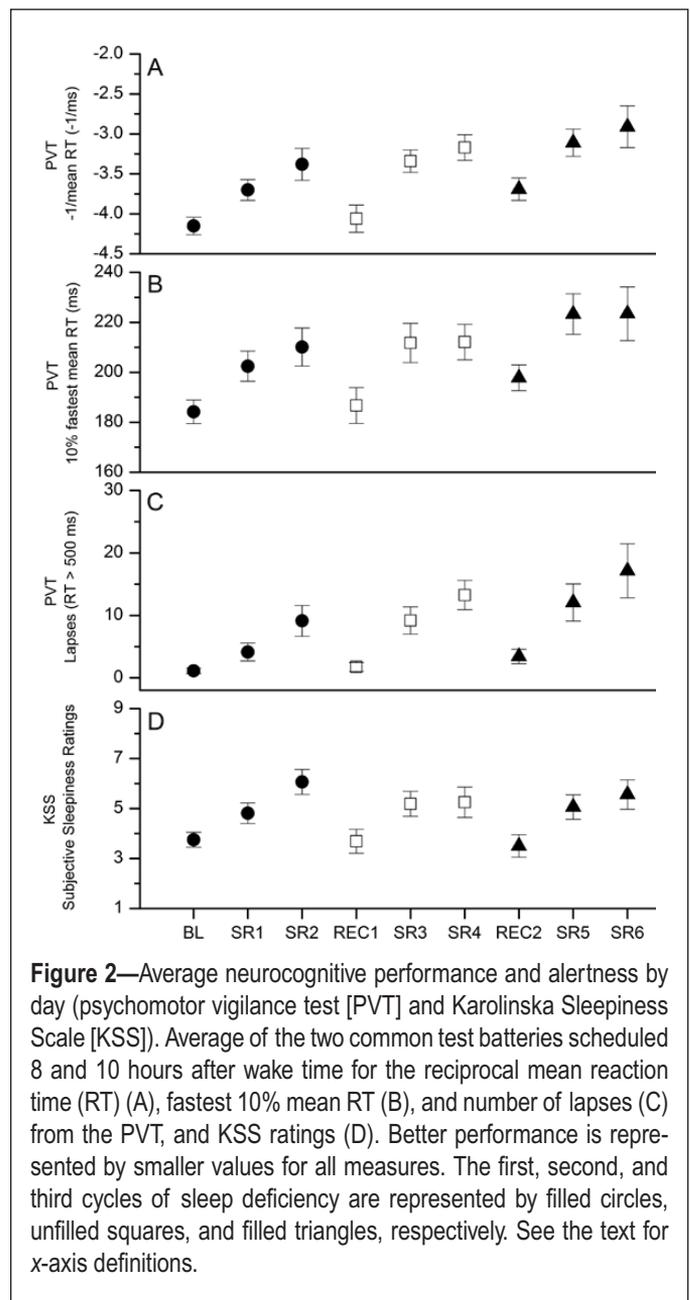

Figure 2—Average neurocognitive performance and alertness by day (psychomotor vigilance test [PVT] and Karolinska Sleepiness Scale [KSS]). Average of the two common test batteries scheduled 8 and 10 hours after wake time for the reciprocal mean reaction time (RT) (A), fastest 10% mean RT (B), and number of lapses (C) from the PVT, and KSS ratings (D). Better performance is represented by smaller values for all measures. The first, second, and third cycles of sleep deficiency are represented by filled circles, unfilled squares, and filled triangles, respectively. See the text for x-axis definitions.

following REC2 compared with BL ($p = .036$, $d_z = 0.79$), but failed to reach significance between REC1 and REC2 ($p = .080$, $d_z = 0.58$). A similar pattern was observed for the number of lapses (BL–REC2 $p = .014$, $d_z = 0.60$; REC1–REC2 $p = .059$, $d_z = 0.57$). There was no difference in KSS after the three 10-hour TIB opportunities ($F_{(2,14)} = 0.80$, $p = .468$).

Time Course and Daily Averages

A limitation of using only two time points for the prior analysis is that results may be confounded by daily variations in circadian phase.[25] For illustration purposes, therefore, the daily average and time course for all PVT and KSS test batteries are shown for the reciprocal mean RT, fastest 10% mean RT, number of lapses, and KSS (Figure 3A–D, respectively). The daily averages exclude test batteries completed within the first 3 hours after waketime due to sleep inertia effects,[26,27] but include
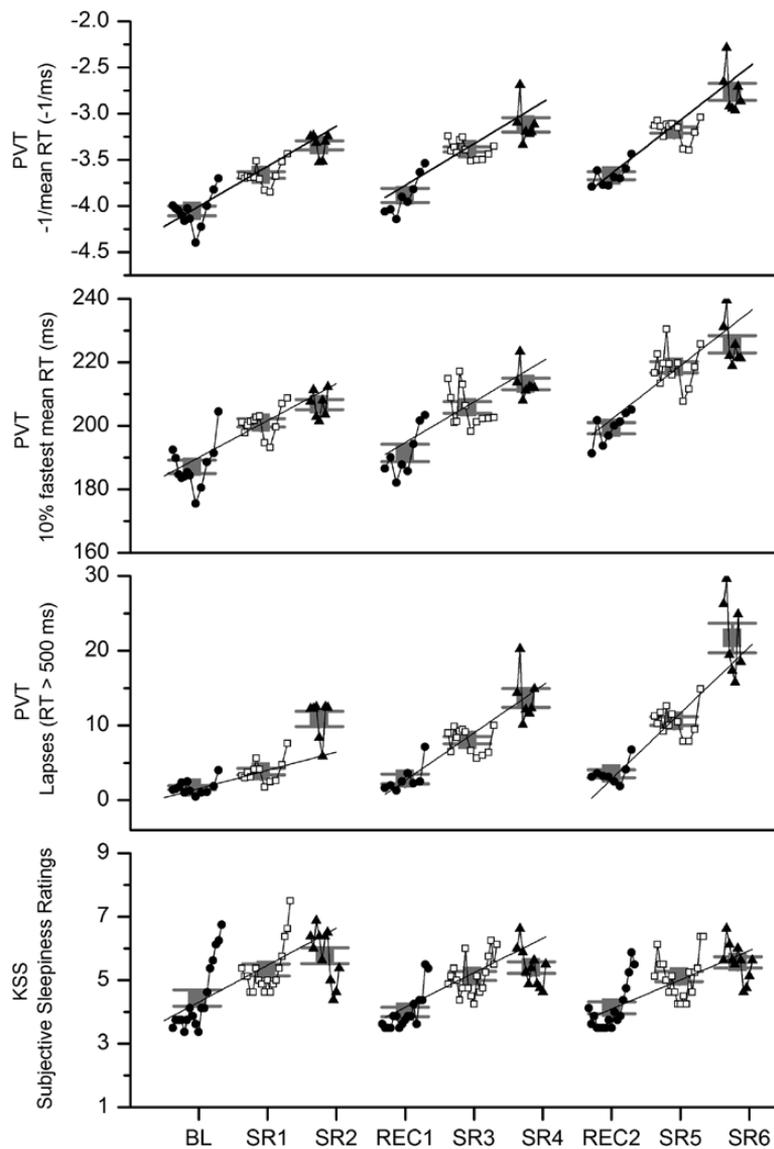
Figure 3—Time course of psychomotor vigilance test (PVT) and Karolinska Sleepiness Scale (KSS) measures. Average time of day pattern for each individual test battery and overall daily average of all test batteries for the reciprocal mean reaction time (RT) (A), fastest 10% mean RT (B), and number of lapses (C) on the PVT and KSS ratings (D). Better performance is represented by smaller values for all measures. Gray symbols with error bars indicate the daily averages. The first, second, and third day of each cycle are represented by filled black circles, open black squares, and filled black triangles, respectively. See the text for *x*-axis definitions.

test batteries completed during light exposure (BL, SR2, SR4, SR6), which had no measurable effect on performance or subjective sleepiness (Supplementary Materials). The statistical results for the daily average were consistent with the results averaged at 8 and 10 hours awake. The main effect of cycle was significant for the reciprocal mean RT ($F(2,14) = 17.85$, $p = .0001$) and all pairwise comparisons were significant (cycle 1–2: $p < .0001$, $d_z = 0.64$; cycle 2–3: $p < .0001$, $d_z = 0.29$; cycle 1–3: $p < .0001$, $d_z = 0.94$). A similar result was observed for the fastest 10% mean RT ($F(2,14) = 10.40$, $p = .0017$; cycle 1–2: $p = .0017$, $d_z = 0.43$; cycle 2–3: $p < .0001$, $d_z = 0.32$; and cycle 1–3: $p < .0001$, $d_z = 0.74$) and the number of lapses ($F(2,14) = 16.43$, $p = .0002$; cycle 1–2: $p = .0017$, $d_z = 0.50$; cycle 2–3: $p < .0001$, $d_z = 0.10$; and cycle 1–3: $p < .0001$, $d_z = 0.57$). There was

no significant main effect of cycle on KSS ($F(2,14) = 0.78$, $p = .4785$) when all test batteries administered more than 3 hours after wake were included in the analysis.

The time courses of PVT and KSS data from each cycle were fit with a linear function to determine whether a change in the rate of impairment across cycles could be observed (Figure 3A–D). The group average time course PVT and KSS data were fit by a simple linear regression weighted by the inverse of the squared standard error in Origin 8.5 Pro (OriginLab Corporation, Northampton, MA) for each cycle. Pairwise *t*-tests were conducted to determine whether the rate of change per day differed across each cycle for each PVT or KSS measure. The rate of change per day in the reciprocal mean RT increased from cycle to cycle (cycle 1 = 0.43 ± 0.23, cycle
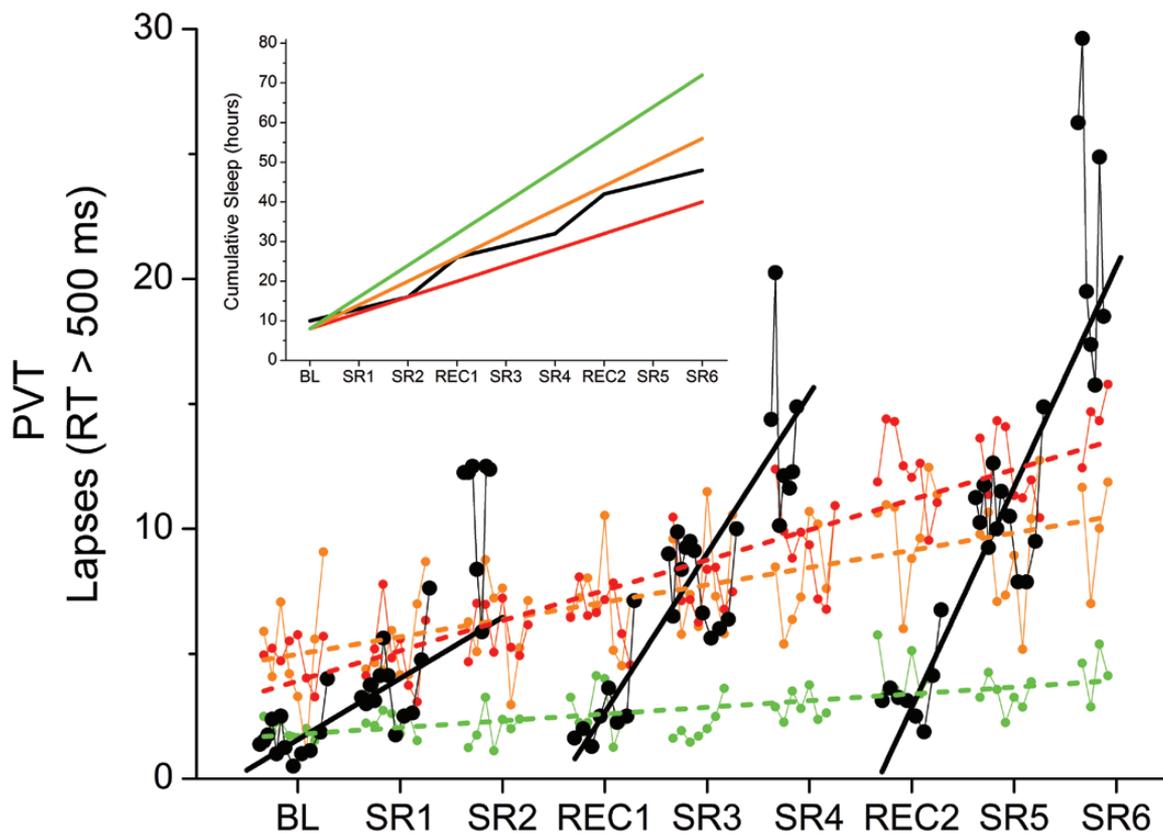
Figure 4—Comparison of PVT lapses between chronic stable and chronic variable sleep deficiency. The time course of the PVT lapses observed during the present chronic variable sleep deficiency (CVSD) study (black filled circles) are plotted against data from an independent control dataset in which participants were scheduled to 4, 6, or 8 hours time-in-bed (TIB) sleep opportunities (red, orange, and green circles, respectively). The solid black line represents the best-fit regression line through each cycle of CVSD, as in Figure 3C, and the dashed red, orange, and green lines represent the best-fit regression lines through the 4-, 6-, and 8-hour TIB sleep opportunity data, respectively. Better performance is represented by smaller values. See the text for *x*-axis definitions. The inset shows the cumulative sleep amounts for the CVSD study (black lines) compared to the 4-, 6-, and 8-hour TIB sleep opportunity data (red, orange, and green lines, respectively).

2 = 0.45 ± 0.40, cycle 3 = 0.58 ± 0.35) and were significantly different from zero (all $p < .03$), but did not differ between cycles (all $p > .05$). Similar results were observed for the fastest 10% mean RT (cycle 1 = 11.6 ± 0.9.3, cycle 2 = 12.8 ± 14.6, cycle 3 = 16.9 ± 10.7). For the number of PVT lapses, the rate of change per day increased from cycle to cycle (cycle 1 = 2.4 ± 2.4, cycle 2 = 6.4 ± 4.6, cycle 3 = 8.8 ± 6.4) such that the rate of change in the number of lapses across the cycle was significantly different between cycle 1 and cycle 2 ($p = .047$) and cycle 1 and cycle 3 ($p = .033$), but not significantly different between cycle 2 and cycle 3 ($p = .216$). These slopes were also all significantly different from zero (all $p < .03$). For the KSS, the rate of change per day decreased from cycle to cycle (cycle 1 = 1.2 ± 1.2, cycle 2 = 1.1 ± 1.1, cycle 3 = 0.92 ± 0.93) and were significantly different from zero (all $p < .02$), but did not differ between cycles (all $p > .05$).

Previous studies under a variety of laboratory conditions have established that when individuals obtain 8 hours or more TIB per night, PVT lapses are consistently low and either slowly increasing or stable, whereas less than 8 hours TIB leads to significant decline in PVT performance, with a greater rate of degradation the shorter the TIB. As a comparison, we plotted data from an independent data set, in which visual PVT lapses were collected from 35 healthy young participants (6 female) who were scheduled to 4-hour ($n = 13$), 6-hour ($n = 13$), or 8-hour ($n = 9$) TIB sleep opportunities for 14 consecutive days under controlled laboratory conditions in a different laboratory environment.[6] Linear increases in PVT lapses were observed with slopes of 0.27 lapses per day ($p < .001$) for the 8-hour TIB opportunity, 0.69 lapses per day ($p < .001$) for the 6-hour TIB opportunity, and 1.21 lapses per day ($p < .001$) for the 4-hour TIB opportunity. The rate of degradation in PVT performance following a 3-hour sleep opportunity in our study is steeper than that of the 4- and 6-hour TIB opportunity on days, which is particularly surprising given that participants on the CVSD protocol had more cumulative sleep than those who participated in the 4-hour TIB protocol. Notably, however, the number of PVT lapses coincided with the 8-hour TIB group on days following a 10-hour sleep opportunity (Figure 4).

Comparison With Mathematical Model Predictions

The population-average PVT lapse data and model simulations for each of the five group-average models are plotted in
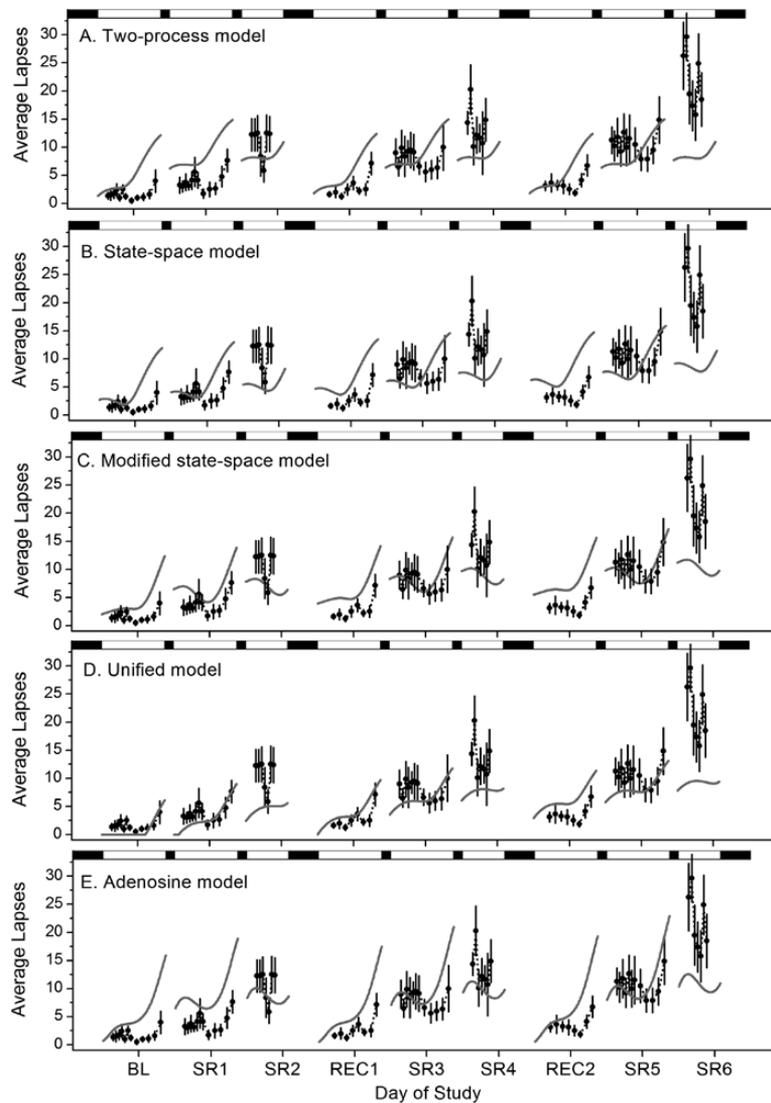
Figure 5—Model predictions of the 11-day chronic variable sleep deficiency (CVSD) protocol. The 11-day CVSD protocol (Figure 1) was simulated using the two-process model (A), the state-space model (B), the modified state-space model (C), the unified model (D), and the adenosine model (E). These model predictions (gray solid lines) were compared to the population-average PVT lapses (filled circles). Sleep and wake episodes (black and white bars, respectively) are plotted across the top of the graph. SR1, SR2, SR3, SR4, SR5, and SR6 represent days following a 3-hour time-in-bed (TIB) opportunity and BL, REC1, and REC2 represent days following a 10-hour TIB opportunity.

Figure 5 and the GoF measures for each day of the protocol are reported in Table 2. Visually, the magnitude of performance impairment during the final day of each cycle, during which subjects were expected to be the most impaired due to a cumulative effect of sleep loss, appeared to be best captured by the adenosine model, which is consistent with the within-day RMSE but not the within-day AIC of the final day of each cycle (Table 2). Overall, the within-day GoF measures provide inconsistent findings. For example, despite the unified model having the lowest within-day AIC values across most study days (where lower AIC reflects a better relative model fit), this model appears to underpredict performance impairment on the final day of each cycle, when impairment was highest. In addition, the models vary with respect to their ability to predict the recovery following each 10-hour TIB opportunity,

as observed in the data. Visually, the change in performance levels from SR2–REC1 to SR4–REC2 appears to be captured best by the adenosine model and the two-process model but not as well by the unified, state-space, and modified state-space models. Again, the within-day GoF measures provide inconsistent findings (Table 2). Overall AIC indicates that predictions of the unified model provide the best GoF, due in part to the lower number of parameters in this model, whereas overall RMSE suggests that predictions of the modified state-space model provide the best GoF. Spearman's rho indicates strong positive correlations between model predictions and data for the modified state-space, unified, and adenosine models (ρ = 0.72, 0.76, and 0.67, respectively; all $p <$ .0001) and a moderate positive correlation between model predictions and data for the two-process model and state-space model

Table 2—Goodness-of-Fit (GoF) Values for Predictions of the Five Group-Average Models to the Population-Average PVT Lapse Data from the 11-day Chronic Variable Sleep Deficiency Protocol.

GoF	Model	Cycle 1			Cycle 2			Cycle 3			Overall
		BL	SR1	SR2	REC1	SR3	SR4	REC2	SR5	SR6	
AIC	Two process	46.40	51.83	**30.31**	34.40	40.88	**37.54**	33.22	42.26	**49.33**	311.58
	State-space	58.33	53.41	52.28	50.94	57.35	53.99	53.30	57.55	63.12	332.67
	Modified state-space	56.43	56.02	50.70	52.09	48.97	53.23	54.93	47.30	64.62	295.94
	Unified	**22.56**	**27.47**	40.43	**24.96**	**36.70**	40.07	**32.62**	**39.72**	50.15	**289.65**
	Adenosine	61.08	65.27	42.32	50.19	56.08	47.49	51.83	51.44	60.14	306.13
RMSE	Two process	4.19	5.26	3.70	4.06	3.33	6.20	3.77	3.53	14.39	5.83
	State-space	3.85	3.13	6.54	4.75	3.69	7.38	5.51	3.72	14.17	6.07
	Modified state-space	3.01	2.96	4.39	3.98	**2.20**	5.26	4.75	**2.06**	11.86	**4.78**
	Unified	**1.43**	**1.75**	6.61	**1.98**	2.57	6.44	**3.20**	2.92	13.22	5.06
	Adenosine	4.31	5.14	**3.21**	4.53	3.50	**4.64**	5.03	2.89	**11.45**	5.20

AIC = Akaike's information criterion; RMSE = root mean square error. Bold values indicate best-fit model for each day for each GoF measure.

($\rho = 0.47$ and 0.39, $p < .0001$). In contrast, the adjusted R^2 is highest for the state-space model (adjusted $R^2 = 0.98$) compared to the other models (two-process, adjusted $R^2 = 0.91$; modified state-space, adjusted $R^2 = 0.53$; unified, adjusted $R^2 = 0.66$; and adenosine, adjusted $R^2 = 0.68$). Taken together, visual, and quantitative GoF indicate that no existing model is capable "out-of-the-box" of predicting both the significant build up in impairment and recovery observed in the data.

As an additional comparison, the predictions of each model were plotted relative to individual PVT lapse data (Supplementary Figure 4 and Supplementary Table 3). As observed from the data, there was significant interindividual variability in the number of lapses across the CVSD protocol, which none of the models could address. Overall, all five models underpredict the number of lapses observed across the protocol, especially during cycle 3.

DISCUSSION

This study demonstrates that CVSD consisting of three cycles of two 3-hour TIB opportunities and one 10-hour TIB opportunity significantly impairs neurocognitive performance. Despite apparent near-complete recovery following each 10-hour TIB opportunity, when rechallenged with chronic sleep loss, PVT performance did not decline from the original BL or with the same trajectory that was observed prior to the recovery sleep. Instead, PVT performance demonstrated a cumulative impairment, suggesting that the apparent recovery was temporary and incomplete. This nonuniform progressive deterioration in PVT performance can be seen both within each cycle of sleep deficiency and between subsequent cycles. In contrast, subjective sleepiness increased in a similar manner within each cycle, but did not exhibit a cumulative increase throughout the study despite the continued degradation of objective performance. None of the models reproduced all features present in the data.

Within each variable sleep cycle, there is a clear and immediate deterioration in performance following one or two 3-hour

TIB opportunities but some restoration of performance following a 10-hour recovery night, as demonstrated previously.[3,6,12,15] The novel aspect of the current study, however, is the immediate rechallenge with successive cycles of sleep restriction and recovery, which causes a progressive deterioration in performance and which mimics schedules in the real world. Although the first 10-hour TIB recovery opportunity (REC1) appears to restore performance close to BL (Figure 2), an immediate rechallenge with a 3-hour TIB opportunity results in a performance decrement that does not repeat the pattern of the first cycle starting from the seemingly recovered BL, but rather is impaired to a level closer to that observed from before the recovery sleep, and subsequently worsens further at increasingly faster rates after the second 3-hour TIB opportunity (Figures 2 and 3). Moreover, the rate of performance decline becomes steeper with increasing number of CVSD cycles, also suggesting a cumulative deterioration in performance (Figure 3). These results suggest that Process S is composed of both fast (acute) and slow (chronic) components, a finding that is supported by the adenosine model. A short recovery sleep appears to restore the fast but not the slow component.

These results are consistent with another laboratory-based study in which participants were scheduled to chronic sleep restriction across all circadian phases in a forced desynchrony protocol.[28] In that study, although participants were scheduled to 10 hours of sleep opportunity per 42.85-hour "day," the effect of the recovery sleep was temporary, and performance continued to decline across the experiment. Our results are also consistent with a field-based study in a group of resident physicians repeatedly exposed to 24–30 hours extended duration work shifts during residency training, in which cumulative sleep deficiency over a 21-day period caused a progressive degradation in PVT performance.[10] Our results are also consistent with the concept of "sleep banking" discussed in Rupp et al.[15] in which extending nightly TIB may confer benefits on performance during subsequent sleep restriction.

None of the five group-average mathematical models tested, of which all except the two-process model incorporate a chronic sleep restriction component, were able to simulate the magnitude of decline across cycles, including elevated performance decrements immediately following the second 3-hour TIB opportunity. Although the data indicate that the initial number of PVT lapses is stable across each cycle (ie, full recovery to BL), the extent of the performance decrement within a cycle cumulatively worsens across subsequent cycles. GoF measures did not reveal clear superiority of one model, but visually the adenosine model appears to best capture the magnitude of performance impairment across cycles, particularly on the last day of each cycle when performance is most impaired, which is consistent with the RMSE on the final day of each cycle. The other four models (ie, the two-process, state-space, modified state-space, and unified models) appear to predict that the amount of performance decrement is approximately stable across each of the three cycles, a pattern consistent with the subjective sleepiness ratings but not PVT performance. The adenosine model also appears to be best able to predict the magnitude of recovery observed following 10-hour TIB opportunities, as evidenced by model predictions at the start of wake, in contrast to the other models.

The adenosine model's ability to predict both acute recovery and chronic worsening of performance across cycles is due to its structure. In all of the models tested, Process S includes both acute (fast) and chronic (slow) subprocesses. In the two-process–based models, the slow subprocess manifests as changes to the homeostatic thresholds of the fast subprocess, and cognitive performance is represented by the fast subprocess. In the state-space model, the upper threshold increases with chronic sleep restriction, meaning wake episodes more rapidly degrade cognitive performance. In the unified model, the lower threshold increases with chronic sleep restriction, meaning sleep episodes less effectively restore cognitive performance. These threshold-based models have difficulty accounting for both the sudden recovery of performance in one night and the progressive worsening of performance across cycles, capturing one phenomenon or the other. In the adenosine model, the fast subprocess is adenosine concentration and the slow subprocess is A_1 receptor concentration. The effect of these two subprocesses on cognitive performance is interactive, due to binding of adenosine at its receptors. In the model, chronic sleep restriction raises concentrations of adenosine (acutely) and A_1 receptors (chronically), leading to worsened performance on both timescales. A single night of recovery restores adenosine concentration to lower levels, resulting in temporarily improved performance, but A_1 receptor concentration remains elevated, amplifying performance impairments when adenosine concentration is rapidly increased again with a subsequent challenge.

Our out-of-the-box challenge of these state-of-the-art group-average mathematical models revealed significant discrepancies between model predictions and experimental data, which suggests that existing parameter estimates of these models may be overfit to their validation datasets and highlights a need for further refinements of all existing group-average mathematical models to account for the unique dynamics of CVSD. Most model developments to date have focused on fitting performance decrements during either acute sleep deprivation or chronic stable sleep restriction. Chronic variable sleep loss, incorporating recovery processes, has received less attention. Furthermore, in the existing group-average models, the circadian system has been assumed to remain stable across chronic sleep restriction, which is unlikely given that variable wake time results in variable light exposure that would be expected to shift circadian phase. In addition, none of the group-average models tested account for direct alerting effects of light on performance. Participants were exposed to relatively bright light levels during wake (~64 µW/cm² [~200 lux] for the first 14 hours of wake and ~23 µW/cm² [~89 lux] during the final 7 hours of wake), which are within range of the light level at which the half-maximum direct alerting effect of light on subjective alertness is achieved (~94.8 lux[29]), as well as brighter experimental light exposures (Supplementary Materials). It would be expected that a model lacking a direct alerting effect of light would overpredict the level of PVT performance impairment under such relatively bright light conditions. The opposite was observed, however, in that all the models underpredicted the level of performance impairment despite the relatively bright background light used throughout the protocol. Future experiments and model development, therefore, should focus on the interaction between the circadian and homeostatic drives under CVSD as well as the effects of direct alerting effects of light on PVT performance during CVSD.

The lack of consistency between objective performance and subjective sleepiness observed in our current study has been previously shown for chronic *stable* sleep deficiency.[6] While subjective ratings of sleepiness during acute sleep deprivation studies (eg, 24 hours or more) are comparable with performance,[6] chronic sleep deficiency appears to pose a more difficult challenge, possibly due to different neurobiological processes between acute and chronic sleep loss, or differences in expectations of sleepiness based on knowledge of time awake. The finding that participants are not subjectively aware of the severity of their performance impairment has important implications for the reliability of self-assessments of sleepiness in operational settings, particularly when chronically but variably sleep deprived.

The rapid relapse of performance after apparent recovery and the residual impairment, which is exacerbated by rechallenging the homeostatic system with sleep restriction, may be due to a change in the sensitivity of homeostatic recovery processes. A substantial amount of SWS was obtained across each cycle, equivalent on average to BL (ie, total SWS of ~255 and ~271 minutes following SR1, SR2, and REC1 and SR3, SR4, and REC2, respectively, which is greater on average per night than the ~81 minutes obtained during the BL 10-hour TIB opportunity). If the total SWS over 3 days of cycle 1 had fully restored homeostatic sleep pressure, then performance on cycle 2 should have followed the same pattern as cycle 1. Instead, performance deteriorated at a higher rate when rechallenged during cycle 2, despite equivalent SWS, suggesting an altered recovery process. Consistent with this hypothesis, McCauley et al.[19] proposed a shift in the physiologic balance of the homeostatic system due to the upregulation and downregulation of adenosine receptors during wakefulness and sleep, respectively, as the system attempts to achieve a new equilibrium as wake is extended and sleep is truncated. This is explicitly represented in the adenosine model. Other studies on the effects of recovery sleep also propose that the physiological

mechanisms underlying the slow recovery following chronic sleep restriction may be due to long-term neuromodulatory changes in the brain.[3,6,15] Our data suggest, however, that these processes can be modulated relatively rapidly within several days.

One-third of the initial study participants withdrew from the study due to complaints of extreme sleepiness. This is a much higher withdrawal rate than our protocols that include 50 hours of constant wakefulness, suggesting that CVSD imposes a different physiological burden than acute sleep loss. Furthermore, the levels of performance impairment following CVSD (ie, a peak average of ~30 lapses during cycle 3) exceed levels observed during chronic stable sleep deficiency (eg, ~17 lapses after 7 consecutive days of 3-hour TIB[3] and ~16 lapses after 14 consecutive days of 4-hour TIB[6]) or acute total sleep deprivation (~16 lapses after 3 days[6]). Comparable levels (~30 lapses) have been observed after a 50-hour constant routine, however.[30] The CVSD schedule of the current study, therefore, may pose a more severe challenge to the homeostatic system than previous chronic sleep deficiency protocols.

Prior studies of up to seven 8- or 10-hour recovery sleep episodes following 5–7 days of chronic stable sleep deficiency[3,12,14,15] may have overestimated the extent or permanence of the recovery by not rechallenging the individuals, given the results of our current study. If true, the overestimation of the benefits of one night of recovery sleep is alarming when considering the consequences for operational settings, especially given that participants in our study were exposed to only two nights of sleep restriction per cycle, which is not an unrealistic or excessive example of real-world sleep patterns. Many industries use self-report or unsophisticated rules to assess fatigue and fitness for duty and recommend sufficient recovery times, but little data exist to inform individuals or companies how much recovery sleep per night over how many days is needed to fully restore and sustain performance to BL levels following chronic sleep deficiency. Both ours and other data provide strong evidence that "sleeping it off" with one night of extended sleep will not resolve performance decrements induced by sleep loss, particularly if a pattern of CVSD persists. The lack of objective data on recovery requirements raises enormous cause for concern when attempting to establish safe working practices or reduce the risk of sleepiness-related accidents and injuries. Notwithstanding simple arithmetic (eg, losing 2 hours sleep per night from Monday to Friday would require an unrealistic additional 5 hours of sleep per night on Saturday and Sunday), commercially available fatigue management programs and state-of-the-art group-average mathematical models based on laboratory data do not estimate accurately the cumulative performance degradation due to chronic sleep deficiency, and therefore overestimate the benefits of recovery sleep on performance. Relying solely on subjective ratings of sleepiness is also a flawed approach, given the disconnect between objective performance and subjective sleepiness that has been observed in ours and other studies.

Our analysis shows that, while the newer group-average models are better than the simple two-process model, most still lack the ability to model the effects of complex sleep cycles on performance accurately. In addition, although the newly proposed adenosine model was able to accurately predict the pattern of performance impairment observed under the CVSD schedule, a comparison of the model predictions to the data at an absolute level suggests significant room for further improvement, especially if the model is to be used to make recommendations on an individual basis. Although additional techniques have been applied to both the two-process model[31] and unified model[32] to generate more accurate individual-level predictions of performance impairment, these techniques assume trait-like differences in response to sleep loss that persist over time, require significant a priori data collection within an individual, and have not been tested outside of laboratory settings. Collectively, these findings should raise concerns about the utility and safety of using predictive models to design and monitor the safety of individual work schedules in real-world settings.

One limitation of the current study is the lack of a well-rested control group with 10 hours TIB opportunities across the 11-day protocol. Our comparison with an historical independent control group with 8 hours TIB sleep opportunities demonstrates that the change in performance impairment in response to subsequent challenges with 3 hours TIB opportunities over 11 days are much greater than those observed in response to 8 hours TIB opportunities over a similar study duration. A similar study in which a well-rested control group slept for 9 hours per night over 7 days[3] also showed minimal increases in performance impairment across the duration of study (eg, less than five lapse difference between BL and the final day of the study), and these effects are also modeled well by existing models of sleep loss.[19,21,33] Although a within-subject control may provide novel information about interindividual vulnerability to CVSD, based on the historical data with 8 hours TIB or 9 hours TIB, we predict that a 10-hour sleep control group would show essentially constant performance and not demonstrate any dynamic change in performance following recovery and rechallenge as observed herein. Based on these prior results, we are confident that the magnitude of performance impairment observed within each cycle and across cycles is due to inadequate dissipation of the accumulated sleep pressure across the study rather than to motivational or laboratory environmental factors. A second limitation of the current study is that only male participants were studied under the CVSD schedule. Female participants were not included in this study to minimize the effects of menstrual cycle phase on sleep and performance measures.[34] Future studies will be needed to investigate the impact of CVSD in women. A third limitation is that participants were restricted from caffeine use both during and prior to the study. This restriction may limit generalizability of our results to real-world settings, as caffeine has been shown to reduce wake-dependent impairment observed in PVT lapses,[35] and individuals exposed to sleep under real-world conditions may use caffeine to mitigate such performance impairment.

In summary, this is the first report of the effects of CVSD on performance and subjective sleepiness. The present CVSD protocol is an extreme sleep–wake pattern characterized by wide variations in daily TST, resulting in severe deficits in alertness and performance. Although such an extreme day-to-day variation in sleep may be more representative of the patterns experienced under operational settings across several safety-sensitive occupations (eg, Lockley et al.[9]; Barger et al.[36]) than the stable chronic sleep restriction schedules that have been previously studied (eg,

Van Dongen et al.[6]; Belenky et al.[3]), few individuals are likely to maintain such a sleep–wake pattern for more than a few days due to the severe impairment encountered. Nevertheless, because existing group-average mathematical models have been developed exclusively on data from studies involving acute sleep deprivation and stable chronic sleep restriction (and/or sleep extension) schedules, the present findings highlight the importance of including performance data from more variable sleep–wake patterns in future model developments. This study therefore has important implications for our understanding of how sleep regulates performance and subjective sleepiness and the time course of the build-up and recovery of neurocognitive impairment under variable sleep patterns. This study also has major implications for occupational fatigue management policies and should prompt a re-evaluation of current modeling approaches, as they are likely to underestimate the degree of performance impairment caused by CVSD in the real world.

REFERENCES

1. National Sleep Foundation. Annual Sleep in America Poll Exploring Connections with Communications Technology Use and Sleep. Washington, DC: National Sleep Foundation; 2011.
2. Groeger JA, Zijlstra FRH, Dijk DJ. Sleep quantity, sleep difficulties and their perceived consequences in a representative sample of some two thousand British adults. J Sleep Res. 2004; 13(4): 359–371.
3. Belenky G, Wesensten NJ, Thorne DR, et al. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. J Sleep Res. 2003; 12(1): 1–12.
4. Borbély AA. A two process model of sleep regulation. Hum Neurobiol. 1982; 1(3): 195–204.
5. Dijk DJ, Hayes B, Czeisler CA. Dynamics of electroencephalographic sleep spindles and slow wave activity in men: effect of sleep deprivation. Brain Res. 1993; 626(1–2): 190–199.
6. Van Dongen HP, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep. 2003; 26(2): 117–126.
7. Sasaki M, Kurosaki Y, Mori A, Endo S. Patterns of sleep-wakefulness before and after transmeridian flight in commercial airline pilots. Aviat Space Environ Med. 1986; 57(12 pt 2): B29–B42.
8. Monk TH, Buysse DJ, Rose LR, Hall JA, Kupfer DJ. The sleep of healthy people—a diary study. Chronobiol Int. 2000; 17(1): 49–60.
9. Lockley SW, Cronin JW, Evans EE, et al. Effect of reducing interns' weekly work hours on sleep and attentional failures. N Engl J Med. 2004; 351(18): 1829–1837.
10. Anderson C, Sullivan JP, Flynn-Evans EE, Cade BE, Czeisler CA, Lockley SW. Deterioration of neurobehavioral performance in resident physicians during repeated exposure to extended duration work shifts. Sleep. 2012; 35(8): 1137–1146.
11. National Sleep Foundation. Summary of Findings: 2005 Sleep in America Poll. Washington, DC: National Sleep Foundation; 2005.
12. Banks S, Van Dongen HP, Maislin G, Dinges DF. Neurobehavioral dynamics following chronic sleep restriction: dose-response effects of one night for recovery. Sleep. 2010; 33(8): 1013–1026.
13. Lamond N, Jay SM, Dorrian J, Ferguson SA, Jones C, Dawson D. The dynamics of neurobehavioural recovery following sleep loss. J Sleep Res. 2007; 16(1): 33–41.
14. Axelsson J, Kecklund G, Akerstedt T, Donofrio P, Lekander M, Ingre M. Sleepiness and performance in response to repeated sleep restriction and subsequent recovery during semi-laboratory conditions. Chronobiol Int. 2008; 25(2): 297–308.
15. Rupp TL, Wesensten NJ, Bliese PD, Balkin TJ. Banking sleep: realization of benefits during subsequent sleep restriction and recovery. Sleep. 2009; 32(3): 311–321.
16. Mallis MM, Mejdal S, Nguyen TT, Dinges DF. Summary of the key features of seven biomathematical models of human fatigue and performance. Aviat Space Environ Med. 2004; 75(3 suppl): A4–A14.
17. Van Dongen HPA. Comparison of mathematical model predictions to experimental data of fatigue and performance. Aviat Space Environ Med. 2004; 75(3 suppl): A15–A36.
18. Klerman EB, St Hilaire M. On mathematical modeling of circadian rhythms, performance, and alertness. J Biol Rhythms. 2007; 22(2): 91–102.
19. McCauley P, Kalachev LV, Smith AD, Belenky G, Dinges DF, Van Dongen HP. A new mathematical model for the homeostatic effects of sleep loss on neurobehavioral performance. J Theor Biol. 2009; 256(2): 227–239.
20. McCauley P, Kalachev LV, Mollicone DJ, Banks S, Dinges DF, Van Dongen HP. Dynamic circadian modulation in a biomathematical model for the effects of sleep and sleep loss on waking neurobehavioral performance. Sleep. 2013; 36(12): 1987–1997.
21. Rajdev P, Thorsley D, Rajaraman S, et al. A unified mathematical model to quantify performance impairment for both chronic sleep restriction and total sleep deprivation. J Theor Biol. 2013; 331: 66–77.
22. Cajochen C. Alerting effects of light. Sleep Med Rev. 2007; 11(6): 453–464.
23. Rechtschaffen A, Kales A. A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Washington, DC: US Government Printing Office; 1968.
24. Borbély AA, Achermann P. Sleep homeostasis and models of sleep regulation. J Biol Rhythms. 1999; 14(6): 557–568.
25. Lockley SW, Dijk DJ, Kosti O, Skene DJ, Arendt J. Alertness, mood and performance rhythm disturbances associated with circadian sleep disorders in the blind. J Sleep Res. 2008; 17(2): 207–216.
26. Jewett ME, Wyatt JK, Ritz-De Cecco A, Khalsa SB, Dijk DJ, Czeisler CA. Time course of sleep inertia dissipation in human performance and alertness. J Sleep Res. 1999; 8(1): 1–8.
27. Wertz AT, Ronda JM, Czeisler CA, Wright KP Jr. Effects of sleep inertia on cognition. JAMA. 2006; 295(2): 163–164.
28. Cohen DA, Wang W, Wyatt JK, et al. Uncovering residual effects of chronic sleep loss on human performance. Sci Transl Med. 2010; 2(14): 14ra3.
29. Cajochen C, Zeitzer JM, Czeisler CA, Dijk DJ. Dose-response relationship for light intensity and ocular and electroencephalographic correlates of human alertness. Behav Brain Res. 2000; 115(1): 75–83.
30. Shekleton JA, Rajaratnam SM, Gooley JJ, Van Reen E, Czeisler CA, Lockley SW. Improved neurobehavioral performance during the wake maintenance zone. J Clin Sleep Med. 2013; 9(4): 353–362.
31. Van Dongen HP, Mott CG, Huang JK, Mollicone DJ, McKenzie FD, Dinges DF. Optimization of biomathematical model predictions for cognitive performance impairment in individuals: accounting for unknown traits and uncertain states in homeostatic and circadian processes. Sleep. 2007; 30(9): 1129–1143.
32. Ramakrishnan S, Lu W, Laxminarayan S, et al. Can a mathematical model predict an individual's trait-like response to both total and partial sleep loss? J Sleep Res. 2015; 24(3): 262–269.
33. Avinash D, Crudele CP, Amin DD, Robinson BM, Dinges DF, Van Dongen HPA. Parameter estimation for a biomathematical model of psychomotor vigilance performance under laboratory conditions of chronic sleep restriction. NSWO. 2005; 16: 39–42.
34. Wright KP Jr, Badia P. Effects of menstrual cycle phase and oral contraceptives on alertness, cognitive performance, and circadian rhythms during sleep deprivation. Behav Brain Res. 1999; 103(2): 185–194.
35. Wyatt JK, Cajochen C, Ritz-De Cecco A, Czeisler CA, Dijk DJ. Low-dose repeated caffeine administration for circadian-phase-dependent performance degradation during extended wakefulness. Sleep. 2004; 27(3): 374–381.
36. Barger LK, Lockley SW, Rajaratnam SM, Landrigan CP. Neurobehavioral, health, and safety consequences associated with shift work in safety-sensitive professions. Curr Neurol Neurosci Rep. 2009; 9(2): 155–164.

SUPPLEMENTARY MATERIAL
Supplementary Material is available at *SLEEP* online.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication February, 2016
Submitted in final revised form September, 2016
Accepted for publication September, 2016
Address correspondence to: Steven W. Lockley, PhD, Division of Sleep and Circadian Disorders, Departments of Medicine and Neurology, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, MA 02115, USA. Telephone: +1-617-732-4977; Fax: +1-617-732-4015; Email: slockley@hms.harvard.edu